

Herausforderungen der Einheitlichkeit von Bahndaten in der Europäischen Union

The challenges of unifying railway data in the EU

Alessa Eckert | Svenja Hainz

Zur gesamtheitlichen Bewertung eines Eisenbahnsystems werden viele hundert Parameter benötigt. Um die Anforderungen an ein „typisch“ europäisches Szenario zu erfüllen, müssen die Daten verschiedener europäischer Eisenbahnsysteme vereinheitlicht werden. Im Shift2Rail-Projekt IMPACT-2 wurde ein Prozess zur einheitlichen Datenbeschaffung entwickelt. Die daraus generierten Ergebnisse können für viele europäische Bahnszenarien extrapoliert werden.

1 Einführung

1.1 Shift2Rail Projektbezug

Die Schiene gilt als Hoffnungsträger hin zu einer nachhaltigen Mobilität. Um dieser Rolle gerecht zu werden, ist es jedoch unabdingbar, dass sich die Bahn weiter hin zu einem zuverlässigen, wirtschaftlichen und kundenfreundlichen Verkehrsträger entwickelt. Das Joint Undertaking Shift2Rail (S2R), eine europäische Initiative, fördert Forschungsprojekte in diversen Bereichen des Bahnsektors durch sein umfassendes Programm. In fünf Innovationsprogrammen

- Rollendes Material
- Infrastruktur
- Leit- und Sicherungstechnik
- IT-Lösungen
- Güterverkehr

werden technische Lösungen zu dessen Weiterentwicklung erarbeitet. In Querschnittsprojekten (Cross Cutting Activities – CCA) werden diese technischen Innovationen zusätzlich hinsichtlich ihres Einflusses zur Zielerfüllung von S2R bewertet [1]. Dazu zählen die Reduktion der Lebenszykluskosten (Life Cycle Costs – LCC) um 50 %, eine Verdopplung der Kapazität sowie eine Verbesserung der Zuverlässigkeit und Pünktlichkeit um 50 % [2]. Um den gesamthaften Zielerreichungsgrad aller Innovationsprojekte zu messen, wird in den Querschnittsprojekten IMPACT-1 und IMPACT-2 ein KPI (Key Performance Indicators)-Modell entwickelt.

1.2 Darstellung des KPI-Modells

Für jeden der drei KPI

- LCC,
- Kapazität und
- Pünktlichkeit

wird ein eigenständiges Modell entwickelt, das die Zielerreichung der S2R-Innovationen in den vier Marktsegmenten

- Hochgeschwindigkeitsverkehr (HGV),
- Regionalverkehr,
- Nahverkehr (Metro) und
- Güterverkehr

ermittelt [3].

Die drei KPI werden hierbei jedoch nicht untereinander optimiert, sondern das Verbesserungspotenzial jedes KPI wird unter zuvor

Several hundred parameters are needed for a comprehensive assessment of the railway system. In the Shift2Rail project IMPACT-2 a process has been developed to ensure that these data can be collected in a coherent manner. To fulfil the requirement of being applied to “typical” European scenarios, data from different European railway systems need to be unified. Fulfilment of this challenging task is producing results which can be extrapolated for many European railway scenarios.

1 Introduction

1.1 Scope of Shift2Rail

In terms of sustainable mobility, railway is considered one of the most promising modes of transport, but it will only be able to keep this promise if the railway is further developed into a cost-efficient, reliable system, which fulfils the expectations of its customers.

The Shift2Rail Joint Undertaking (S2R) is a European initiative funding rail research projects to develop solutions for the railway system based on its comprehensive programme. It consists of five Innovation Programmes

- Rolling Stock
- Infrastructure
- Command, Control and Signalling
- IT Solutions
- Freight rail

in which technical solutions are being developed to enhance functions of the railway system. To compliment these technological innovations, the Cross-Cutting-Activities (CCA) have been identified which are transversal subjects and relevant for many subsystem [1]. Some of the projects launched under CCA aim to estimate the influence of the S2R innovations towards the objectives of the S2R initiative. Goals are among others the reduction of Life-Cycle-Cost (LCC) by 50 %, the doubling of Capacity and the enhancement of Reliability and Punctuality by 50 % [2]. To assess the degree of target achievement a KPI model is being developed within the S2R projects IMPACT-1 and IMPACT-2.

1.2 Description of the KPI model

For each of the three KPI

- LCC
- Capacity and
- Punctuality

a model structure is set up to assess the potential of S2R innovations in 4 different market segments

- high speed passenger rail
- regional passenger rail
- urban passenger rail (metro) and
- rail freight [3].

definierten Kriterien und unter Konstanthaltung der anderen beiden KPI optimiert.

Das grundlegende Vorgehen ist für alle drei Modelle und vier Marktsegmente gleich. Zunächst wird ein Ist-Szenario definiert und es werden die zur Beschreibung notwendigen Parameter bestimmt, z. B. die Investitionskosten für Bremsen oder die Verspätungsminuten, ausgelöst durch Weichenfehler. Diese Ist-Situation ist festgelegt als ein eingeschwungener Zustand bezogen auf das Jahr 2013 und soll Allgemeingültigkeit für Europa besitzen [4]. Diese Struktur ermöglicht es, Parameter im Modell zu integrieren, die durch die S2R-Innovationen beeinflusst werden, und sie zu einem Ergebnis im jeweiligen Marktsegment zu aggregieren. Im Anschluss können die einzelnen Auswirkungen der Innovationstechnologien auf die entsprechenden Parameter angewendet und ein Zukunftsszenario mit allen S2R-Innovationen abgebildet werden. Hieraus werden dann die Auswirkungen auf die Zielerreichung der KPI LCC, Kapazität und Pünktlichkeit berechnet [5].

Zur Umsetzung dieser Methodik und Bestimmung der Ist-Szenarien aller vier Marktsegmente werden über 2000 einzelne Parameter benötigt. Im Folgenden werden daher die Herausforderungen der Beschaffung dieser Daten sowie ihrer Nutzung unter der Prämisse, dass die Ist-Szenarien europaweit anwendbar sind und ein akzeptabler Grad von Kohärenz erreicht wird, beschrieben.

2 Herausforderungen der Beschaffung von Bahndaten

2.1 Anforderungen an Daten zu einer Bewertung auf europäischer Ebene

Um das Bahnsystem der Zukunft mit dem heutigen Bahnsystem vergleichen zu können, müssen die Verbesserungen durch die S2R-Innovationen gegen einen Ist-Zustand gemessen werden. Da sich die Forschungsaktivitäten von S2R auf die Europäische Union beziehen, hat das Ist-Szenario zusätzlich den Anspruch, für alle EU-Mitgliedsstaaten anwendbar zu sein. Für die Entwicklung der Ist-Szenarien bedeutet dies, dass eine „typische“ europäische Infrastruktur und Züge zugrunde gelegt werden müssen.

Durch die dafür notwendige Zusammenführung von Daten aus unterschiedlichen europäischen Ländern, ergeben sich verschiedene Herausforderungen. Einer der wichtigsten Aspekte ist die Vergleichbarkeit der Daten. Einheitliche Begriffsdefinitionen sowie Erhebungsmethodiken sind essenziell. Dies sicherzustellen, ist bereits innerhalb eines Landes schwierig und noch herausfordernder in einem europaweiten Kontext.

Historisch bedingt unterscheiden sich Technologien und Betriebsabläufe in Europa stark, weshalb es kein klassisch europäisches Szenario gibt. Das Ziel ist es daher, das Ist-Szenario so zu definieren, dass es keine Länderspezifika enthält und somit für die Mehrheit der Bahnstrecken in Europa anpassbar ist.

Folglich ist die Anwendbarkeit und Güte der Ist-Szenarien abhängig von der Verfügbarkeit von Datenquellen, gleichzeitig bedeutet eine hohe Verfügbarkeit unterschiedlicher Datenquellen auch eine höhere Bandbreite an Begriffsdefinitionen und Erhebungsmethodiken. Hierbei kann es sich um naheliegende Unterschiede, wie die Grenze, ab der ein Zug als verspätet in die Statistik eingeht, handeln oder auch um versteckte Unterschiede, welche tiefergehende Analysen benötigen. Ein Beispiel hierfür sind die Bestandsinvestitionskosten für die Gleisinfrastruktur. Werte dieser Kosten können sich je nach der angewandten Erfassungsmethodik der Netzbetreiber unterscheiden; sie können Kosten für die Gleise, Schwellen, Oberbau, Unterbau oder einen Mix der Bestandteile umfassen sowie für ein Richtungsgleis oder Doppelgleise gelten. Zusätzlich können, je nach Land, Subventionen berücksichtigt sein, sodass

Thereby the three KPI, which in reality are competing, are not optimised against each other, but for every KPI the potential of improvement through Shift2Rail is evaluated under predefined conditions, including a fixing of the other two KPI.

The general principle is the same for the model of all three KPI and all four market segments. First a baseline is defined and the different necessary parameters for the model are determined, e.g. the investment cost for brakes or the delay minutes caused by switch failures. It is defined that this baseline should be set as in operation 2013 and being applicable within Europe [4]. This structure allows it to integrate the parameters influenced by the S2R innovations in the model and merge them to a result for the railway system of the respective market segment. Afterwards the individual effects of the S2R innovations are applied on their respective parameters and the situation in the future, with all S2R innovations implemented, can be calculated and hence the change of LCC, Capacity and Punctuality be evaluated [5].

To apply this methodology for the KPI model and to set up the baseline for all four market segments more than 2.000 parameters are needed. In the following, the challenges to collect this data and ensure its applicability for most parts of Europe while securing a certain degree of coherence will be discussed.

2 Challenges of collecting railway data

2.1 Requirements on data for an impact assessment on European level

In order to be able to compare the railway system as it is today with a future railway system including all the S2R innovations, the improvements that these innovations provide, need to be measured against a to-be-defined baseline. As the research activities in S2R are targeted for the European Union, this baseline has additionally the entitlement to be applicable for every member country. For scenarios in S2R this would mean to have a “typical” European line and train considered.

Consequently, various challenges arise when working with railway data from different countries in the EU. It is important that the collected data is comparable. Common definitions of data items and measurement techniques are therefore essential. This is already a challenge when trying to obtain data from various sources within the same country and provides even more obstacles when collecting railway data across Europe.

For historical reasons technologies and even rules of operation differ widely among Europe. Therefore, a “typical” scenario cannot be found. Thus the goal of the scenario definition here is to create a scenario which is general enough to not include specifics for one country in Europe, but to be adaptable for most countries.

Of course, the grade of adaptability is relying on the data sources available from which the scenarios are built, but at the same time the more data sources are considered within the scenarios, the wider the differences of definitions of data are. This can be more obvious differences such as the delay threshold which has to be exceeded for trains to be counted into the delay statistics. There are however also hidden differences that need more effort in order to be made comparable. One example is track cost. A value for this parameter can -depending on the infrastructure manager's methodology of monitoring costs - contain the cost for the rail, sleepers, ballast, superstructure or any mix of the mentioned parts. It can further refer to one line or double track. And depending on the country, subsidiaries can be included in the cost, which can make the value provided by the respective infrastructure manager appear lower than its actual value. Even though ex-

die Kosten vom Netzbetreiber geringer ausgewiesen werden als es für die Bewertungsmodelle relevant ist.

Trotz eines erhöhten Aufwands bei der Bereinigung dieser Daten ist ein länderübergreifender Vergleich möglich. Es gibt jedoch auch Parameter, die sich grundlegend zwischen den einzelnen Ländern unterscheiden, wie z.B. Parameter zur Beschreibung von Betriebsabläufen, aber auch Infrastrukturdaten. So unterscheidet sich die Spannung der Oberleitung genauso wie die Sicherungslogik und Anlagen zur Gleisfreimeldung. Dies gilt es bei der Vereinheitlichung von Daten zu berücksichtigen, um europäische Durchschnittswerte dort zu vermeiden, wo sie falsche Ergebnisse erzeugen.

Auch innerhalb einzelner Länder sind Daten nicht immer homogen und jede Bahnstrecke hat zusätzliche Spezifika. Unterschiedliche Regionen innerhalb eines Landes können spezielle Eigenheiten bezüglich Wetterverhältnisse und Topografie aufweisen. In Deutschland ist das Eisenbahnsystem z.B. angepasst an flache Landschaft mit starken Windverhältnissen im Norden und Hochgebirge mit hohem Schneefallrisiko im Süden. Ebenso gibt es charakteristische Unterschiede zwischen den Baureihen von Zügen, Lokomotiven und Güterwagen. Alle Einflussfaktoren auf nationaler Ebene und europäischer Ebene müssen bei der Ermittlung von Durchschnittswerten für das Ist-Szenario berücksichtigt werden. Zusätzlich zu diesen unterschiedlichen Definitionen für Daten kommt es ebenfalls vor, dass verschiedene Fachdisziplinen im Eisenbahnwesen gleiche Begriffe für Daten verschieden auslegen. So gibt es insbesondere für Kostenbegriffe unterschiedliche Interpretationen, wie z.B. ob Instandhaltungskosten Personalkosten beinhalten. Auch hier ist daher eine klare Kommunikation mit allen Beteiligten essenziell.

Auch wenn es das Ziel des Projektes IMPACT-2 ist, einheitliche europaweit gültige Daten zu generieren und auszuwerten, ist eine Unterscheidung des Ist-Szenarios in den verschiedenen Marktsegmenten (HGV, Regionalverkehr, Metro, Güterverkehr) aus verschiedenen Gründen mit teils spezifischen Ansätzen sinnvoll:

- Die Betriebsabläufe unterscheiden sich (z.B. Zugdichte, Hauptverkehrszeiten)
- Nicht alle Infrastrukturanlagen, die im KPI-Modell berücksichtigt werden, sind für alle Marktsegmente relevant (z.B. Rangierbahnhöfe und Umschlagsanlagen oder keine Bahnübergänge im HGV)
- Der KPI Pünktlichkeit wird für die verschiedenen Marktsegmente unterschiedlich erhoben (z.B. wird die Grenze, ab der ein Zug als verspätet gilt, im HGV mit 5,5 Minuten angesetzt, im Güterverkehr mit 15,5 Minuten [6])
- Der KPI Kapazität wird im Personen- und im Güterverkehr unterschiedlich definiert (Spitzenstunde versus 24-Stunden Intervall)
- Effekte wie Lärmbelastung werden nicht für alle Marktsegmente erhoben.

Um ein Ist-Szenario für jedes Marktsegment zu definieren, sind unterschiedliche Datensets notwendig. Reale Szenarien haben den Vorteil, dass auf ihnen berechnete Ergebnisse stärker der Realität entsprechen. Aus diesem Grund wurde jedem Marktsegment ein realer Korridor zugrunde gelegt, der im weiteren Prozess mit den erhobenen Daten verallgemeinert wurde. Um sicherzustellen, dass die Ergebnisse des Projekts relevant für viele europäische Länder sind, wurden spezifische Eigenheiten generalisiert und sensible Daten geglättet. Das Vorgehen hierzu wird in Abschnitt 3 erläutert.

2.2 Identifikation relevanter Daten

Die notwendigen Daten zur Ermittlung der Effekte, die die S2R-Innovationen auf die KPI LCC, Kapazität und Pünktlichkeit haben, können in zwei Gruppen gegliedert werden: Eingangsdaten und Progressdaten.

amples like collecting cost data for track requires effort, they can be compared once it is known what is included in the value.

There are however data that differ in their nature, e.g. operational proceedings among railway operators. The power supply from the catenaries and the distribution of substations for example differ between the countries as well as the control system and assets used to ensure train integrity and to monitor the occupancy of the track. This has to be kept in mind when comparing data; finding an average across Europe is not always useful.

And even within one country data is not always homogeneous; every line has its own specifics. Besides this particularity of every single line, in larger countries also weather conditions and topography can differ quite substantially e.g. in Germany with its flat landscape prone to storms in the North and the mountainous landscape with heavy snow during the winter months in the South. Also the different train and locomotive classes have individual characteristics. Both have to be considered when working with average data within one country and even more across Europe.

Additionally to the differences in data definition, it also happens that different disciplines within the railway community have non-identical definitions for the same data. Especially cost categories are often understood differently, e.g. if the cost for staff executing maintenance is included in maintenance cost. Again a good communication is crucial to make sure that provided data is understood correctly.

Even though the objective of the project is to collect and assess data in a coherent way throughout the EU countries, a differentiation between the market segments (high speed, regional, metro and freight rail) with partially individual approaches is useful for various reasons, e.g.:

- Operations are organised differently (e.g. frequency of trains, time of main operation)
- Not all assets that are included within the KPI assessment are relevant in all market segments (e.g. marshalling yards or terminals or no level crossings in high speed traffic)
- The KPI punctuality is measured differently for the individual market segments (e.g. the definition of delayed trains in high speed is by a delay more than 5.5 minutes, for freight rail it is by more than 15.5 minutes [6])
- A different definition for the KPI capacity in passenger market segments and freight transport (peak hour in comparison to full day)
- Operational aspects such as noise pollution are not relevant for all market segments

To define one scenario for each of the market segments, different kinds of data are necessary. As real case scenarios have the advantage that their results are more realistic, the scenarios for the market segments have been derived from real lines. In order to ensure, that the results obtained from the KPI can be extrapolated to any country in the EU, the scenarios have however been generalised by excluding real line specifics and smoothing sensible data. To ensure a suitable generalisation, a process has been established which is further described in section 3.

2.2 Identification of suitable data

The data needed in the KPI model to determine the effect the S2R innovations have on LCC, Punctuality and Capacity can be clustered in two groups: Input data and improvement data.

Input data are used to describe the baseline scenario, so the status of the system as it is today or rather has been around 2013. As there are four market segments, several values for each parameter are possible.

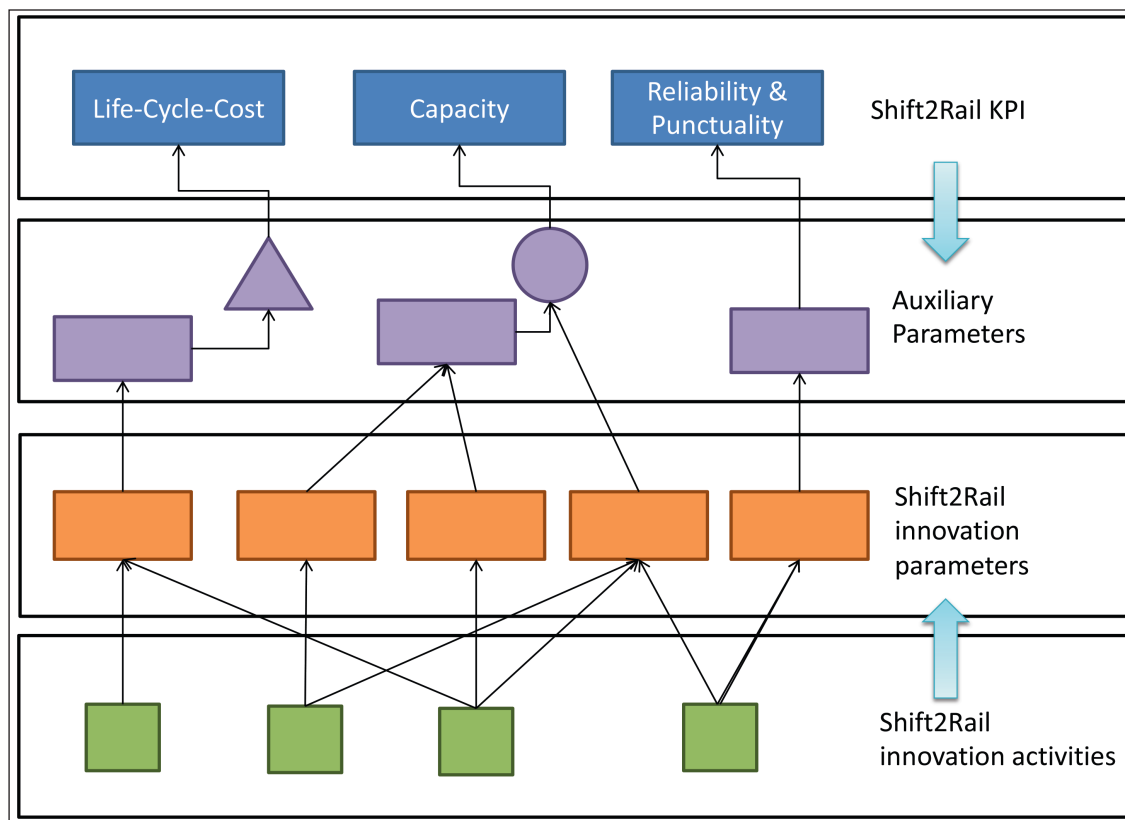


Bild 1: Zweistufiger Ansatz zur Identifikation von Daten

Fig. 1: Twofold approach to data identification

Quelle / Source: DLR

Eingangsdaten werden benötigt, um das Ist-Szenario, also das heutige bzw. das Eisenbahnsystem in 2013, zu beschreiben. Durch die Unterscheidung in vier Marktsegmente kann es mehrere Werte für jeden Parameter geben.

Die Eingangsdaten beschreiben das Ist-Szenario in dem Detailgrad, der notwendig ist, um die Effekte der S2R-Innovationen darzustellen. Zur Quantifizierung dieser Effekte sind Abschätzungen zu dem Einfluss, den die Innovationen auf das Bahnsystem der Zukunft haben, nötig. Diese Abschätzungen werden als Progressdaten bezeichnet und beschreiben die Entwicklung vom Ist- zum Soll-Szenario. Bereitgestellt werden sie direkt aus den Innovationsprojekten heraus. Da der Fokus in diesem Beitrag auf den Daten des Ist-Szenarios liegt, werden sie hier nicht weiter erörtert. Grundsätzlich besteht die Methodik zur Bestimmung der Parameter und Beschaffung der dazugehörigen Daten aus zwei Ansätzen. Zum einen wird ein Top-down-Ansatz ausgehend von den drei zu berechnenden KPI LCC, Kapazität und Pünktlichkeit verwendet. Dieser Ansatz bestimmt die Parameter, die für die Kalkulation der KPI notwendig sind.

Zum anderen werden in einem Bottom-up-Ansatz aus den S2R-Innovationsprojekten Parameter bestimmt, da in den Projekten definiert wird, welche Verbesserungen im Eisenbahnsystem das jeweilige Ziel sind. Die S2R Innovationen geben den Detaillierungsgrad vor, mit dem die einzelnen Parameter in das KPI-Modell eingehen (Bild 1).

Zum Zusammenführen der Daten zu den drei KPI werden Hilfsparameter genutzt, die über den Top-down-Ansatz ermittelt werden. Bei diesen teilweise von der EU empfohlenen Werten kann es sich z.B. um einen Bewertungszeitraum von 30 Jahren für Eisenbahnprojekte oder einen Diskontierungssatz von 3 % handeln. Die Hilfsparameter dienen nicht dazu, die Szenarien zu beschreiben, sondern die erhobenen Daten verschiedener EU-Länder vergleichbar und aufrechenbar zu machen. Dafür wurden die folgenden Einheiten je KPI definiert:

Input data make up the baseline scenario in the granularity that is needed to show the effects of the S2R innovations. In order to quantify the effects of these innovations, estimations on their contribution to the railway system of the future is necessary. These data are referred to as improvement data and describe the transition from the baseline to the future scenario. These data are provided through the projects, developing the S2R innovations, and will not be further discussed in this article.

In general the approach used to define the parameters and therefore data that need to be collected is done twofold. In a first step, a top down approach is used. Coming from the three KPI LCC, Capacity and Punctuality it is assessed which data are needed for their calculation.

In a second step, a bottom up approach from the S2R innovation projects is used, as these define which parameters their improvements aim at. The S2R innovations set the level of detail with which each of the parameters has to feed into the KPI model (see Fig. 1). When aggregating the data from the bottom to the three KPI, auxiliary parameters are used, which were identified through the top down approach. These partially recommended values by the EU can be a common assessment period of 30 years for railway projects or a discount rate of 3 %. The auxiliary parameters are not needed to describe the different scenarios for the market segments per se but to ensure that the data collected from different EU members are comparable and can be aggregated. The units

- euro per person-kilometre or metric ton-kilometre for the LCC model
 - delay minutes per year for the punctuality model and
 - passenger per hour and tons per day respectively for passenger and freight transport in the capacity model
- have been defined [5, 7]. The data needed to calculate these values however are not all in the same format. More scenario specific auxiliary parameters such as the length of the scenario are

- EUR pro Personenkilometer bzw. Tonnenkilometer für LCC,
- Verspätungsminuten pro Jahr für Pünktlichkeit und
- Fahrgäste pro Stunde bzw. Tonnen pro Tag für Kapazität im Personen- und Güterverkehr [5, 7].

Die für diese Berechnungen benötigten Daten werden jedoch nicht alle in der gleichen Einheit erhoben. Zusätzliche Hilfsparameter wie eine definierte Länge des Eisenbahnkorridors im Ist-Szenario müssen daher festgelegt werden, um z.B. Mengen pro Kilometer und pro Gesamtnetz vergleichbar zu machen.

Je aggregierter die zu erhebenden Daten sind, desto höher die Gefahr, dass die Erhebung nicht für alle Parameter in identischer Weise erfolgt ist. Je detaillierter die Daten jedoch sind, desto höher ist der Aufwand der Erhebung und somit die Gefahr, nicht alle für die Kalkulation benötigten Parameter zu erhalten.

3 Datenquellen und ihre Rolle in der Beschaffung einheitlicher Bahndaten

Um eine stabile Datenbasis für die Ist-Szenarien zu erhalten, werden im IMPACT-2-Projekt verschiedene Datenquellen (DQ) genutzt. Einige der Quellen werden zum direkten Dateninput für die Ist-Szenarien herangezogen (primäre DQ), andere werden zur Verifizierung der Daten aus den primären DQ genutzt (sekundäre DQ). Die wichtigsten DQ in IMPACT-2 sind die europäischen Projektpartner, unter ihnen diverse Eisenbahnverkehrsunternehmen (EVU), Eisenbahninfrastrukturunternehmen (EIU), Forschungseinrichtungen und Industriepartner. Sie werden zum einen als primäre DQ einbezogen, indem sie Daten zur Infrastruktur, rollendem Material oder zum Betriebsablauf bereitstellen, zum anderen stellen sie als sekundäre DQ die Kohärenz für Daten aus Drittquellen sicher.

Als primäre DQ werden die gleichen Daten von verschiedenen Partnern aus unterschiedlichen Mitgliedsstaaten gesammelt, um eine möglichst große Bandbreite an Länderspezifika zu erfassen. In einem zweiten Schritt wird diese Sammlung an Daten für einen Parameter von Experten innerhalb des IMPACT-2-Projektes geglättet, sodass der Parameter für die Mehrheit der EU-Länder repräsentativ ist (Bild 2). Die Festlegung des Datenwertes für jeden Parameter erfolgt nach bestimmten Anforderungen.

Das durch die gesammelten Parameter beschriebene Szenario darf nicht exakt mit einer realen Strecke übereinstimmen, da jede reale Strecke besondere Eigenheiten aufweist. Da diese Eigenheiten nicht repräsentativ für andere Korridore sind, würde so die Möglichkeit der Extrapolation von Ergebnissen limitiert werden. Des Weiteren können einzelne Daten, die notwendig sind zur Beschreibung des Ist-Szenarios, hochsensibel sein. Die DQ müssen in diesen Fällen geschützt werden. Die Verwendung von reinen Daten einer realen Strecke könnte jedoch indirekt die Quelle aufdecken, wenn durch diese Daten Rückschlüsse auf den Ort oder den Zugtyp möglich sind. Trotzdem beruhen die verwerteten Daten auf realen Strecken, um sicherzustellen, dass berechnete Ergebnisse so realistisch wie möglich sind und die Parameter eine Ist-Strecke beschreiben, die in verschiedenen Teilen Europas zu finden sein könnte, ohne einem genauen Ort zu entsprechen.

Das Wissen von Experten aus dem Eisenbahnwesen wird zusätzlich genutzt, um Parameter, zu denen nur einzelne oder sehr wenige Daten erhoben werden können, zu verifizieren. Je größer die Anzahl an unabhängigen DQ, die zur Bestätigung eines Parameters vorhanden sind, desto stabiler ist das Ergebnis.

Des Weiteren wurden Open-Source-Daten etwa von ERA (European Railway Agency) oder PRIME (Platform of Railway Infrastructure Managers in Europe) sowie nationalen DQ identifiziert. Daten europäischer Organisationen haben den Vorteil, dass die zur Ver-

needed for example to make asset data per km and asset data per network comparable.

The more aggregated data already are the higher is the chance that the aggregation is not done in the same way in the different countries. The less aggregated the data are however, the more individual numbers have to be collected which increases the effort and consequently the likelihood to not get everything that is needed to perform the assessment.

3 Data sources and their role in obtaining unified railway data

To gain a stable data basis for the baseline scenarios, different sources for data are used within the IMPACT-2 project. Some sources are used to provide data directly to the scenarios (primary sources) whereas others are used to verify the data coming from the primary sources (secondary sources).

In the IMPACT-2 project, the most valuable source are the project partners, among them various railway undertakings (RU), infrastructure managers (IM), research companies and industrial rail partners from across Europe. All of them are either used as primary sources to obtain baseline data on infrastructure, rolling stock or operational parameter for the different EU members or as secondary sources to ensure the coherence between the data from other sources.

As primary sources, the same set of data is collected from partners of different EU members, to ensure that a broad range of country specifics are captured. In a second step a group of experts within the IMPACT-2 project analyses the data and smooths those in a way that they are representative for a majority of the EU countries (see Fig. 2). The determination of each data value is done according to certain requirements.

The scenario that is described by the parameters should not be exactly the same as a real-world scenario as each railway corridor has too many specific characteristics. These are not representative for any other line and would therefore limit the possibility to extrapolate the results. Further, individual data which are needed to describe the corridor can be sensible and the source of the input shall not be disclosed. Providing all data for a real-world corridor could however indirectly reveal the source when experts are able to conclude the location of the corridor or the train type used from the data provided. Nevertheless, the initial data gathered are from different representative real world corridors to ensure that the results calculated are as realistic as possible and the scenario described with the parameters represents a railway corridor that could be found in many places in Europe without being directly linked to one specific region.

The knowledge of experts within the railway business can however also be used to verify data that could only be gathered from very few sources. The higher the number of independent sources confirming the same data, the higher the stability of the results. On top, other sources have been identified such as railway-related open source data from ERA or PRIME or national governmental data sources. Data from European organisations have the advantage that the data provided are often either a European average or data for each EU member. These data sources can again be used to provide baseline data or to verify results. Another example for the verification process is articles on European railway projects. These can contain data on cost and number of assets for a specific project, which can be compared to the IMPACT-2 scenarios for the respective market segment.

If data for certain parameters cannot be obtained directly due to their sensitivity, a two-step approach has been used. In a first

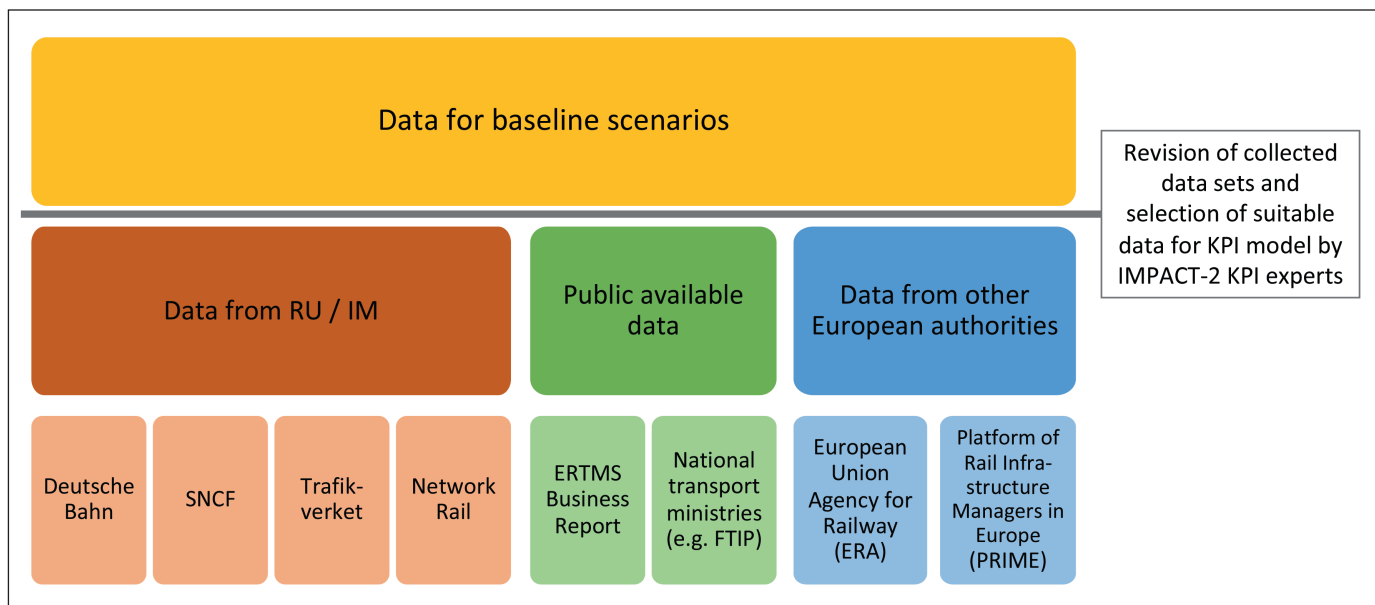


Bild 2: Primäre und sekundäre DQ für Eisenbahndaten in der EU

Fig. 2: Primary and secondary sources of railway data in the EU

Quelle / Source: DLR

fügung gestellten Daten entweder bereits als europäischer Durchschnitt vorliegen oder für jedes EU-Land verfügbar sind. Sie können je nach Bedarf als direkter Input der Ist-Szenarien oder zur Verifizierung bereits erhobener Daten genutzt werden. Zum Verifizieren kommen zusätzlich Berichte über Europäische Bahnprojekte in Betracht. Diese enthalten häufig Angaben zu Anzahl und Kosten verschiedener Bauvorhaben, die mit den Daten aus den jeweiligen IMPACT-2 Ist-Szenarien abgeglichen werden können.

Für Daten mit erhöhter Vertraulichkeit wurde ein zweiteiliger Prozess entwickelt. Zunächst werden von Partnern aus der Forschung Literaturrecherchen durchgeführt. Die so identifizierten Daten werden dann in einem zweiten Schritt Expertengruppen vorgelegt, in denen diskutiert wird, ob die Daten in einer akzeptablen Größenordnung vorliegen.

Für technische Daten, wie den Energieverbrauch, der zu einem Großteil von Zugcharakteristika abhängt und weniger von Länderspezifika, wurden Ergebnisse aus anderen S2R-Projekten herangezogen. Die Eignung solcher Daten muss allerdings aus zwei Blickpunkten betrachtet werden. Die Berechnungen und die Erhebung der Daten geschieht in den S2R-Projekten in einem Detaillierungsgrad und Aufwand, der im Rahmen eines reinen Bewertungsprojektes nicht realisierbar ist. Dadurch sind die Ergebnisse jedoch auch an die spezifischen Annahmen geknüpft, die in den Projekten verwendet werden. Zur Berechnung des Energieverbrauchs werden z. B. Annahmen zu Gewicht, Geschwindigkeit, Beschleunigung und Bremsverhalten getroffen, welche wiederum abhängig sind von der Zugzusammensetzung, dem Triebfahrzeug sowie der gegebenen Infrastruktur. Die Weiterverwendung der Ergebnisse anderer Projekte setzt daher eine genaue Betrachtung aller getroffenen Annahmen voraus. Nur wenn die Rahmenparameter der individuellen Projekte mit denen in IMPACT-2 getroffenen Annahmen übereinstimmen oder angepasst werden können, ist es möglich, die Ergebnisse in der Bewertung zu nutzen.

4 Zusammenfassung und Ausblick

Die Herausforderungen, die die Vereinheitlichung europäischer Bahndaten mit sich bringen, sind insbesondere zurückzuführen

step intensive desktop research has been conducted from the participating research companies on published papers and wider open source data. The data gathered in this process have then in a second step been presented to expert groups to verify in discussion if they are in the correct order of magnitude.

For more technical data such as energy consumption which depend to a high extent on the rolling stock and less on country specifics, results from other S2R projects have been used. Using their data as input or improvement data can be seen twofold. The calculations undertaken within these projects provides data in a very detailed, well researched quality, that otherwise could not be obtained within the scope of an assessment project. On the other hand, the data resulting from these projects have been calculated according to very specific requirements. Especially in the case of energy consumption, the data calculated depend on many input parameters such as weight, speed and acceleration parameters of the train which depend on the locomotive, train composition and infrastructure layout. Therefore using such output data from other projects, requires a detailed comparison of assumptions taken for both projects. Only if the framework parameters of the individual projects can be adapted to the IMPACT-2 scenarios or are calculated in a way that they can be extrapolated, the output data from these projects can be used.

4 The Conclusion and the prospects

The challenges that have been faced when working with data from different sources in Europe are especially related to their definition. Among EU members and even within each country, data is collected in various ways and in some cases even when data from different sources seem to be the same, they can differ because of wider definitions for data of the same name.

Thus, IMPACT-2 has used a process to generate generic scenarios for four market segments (high speed, regional, metro and rail freight) of the railway system by not only using various sources for data but also merging the expertise from different countries and stakeholders of the railway system in Europe.

auf deren Definition. Zwischen EU-Mitgliedern sowie innerhalb der Länder werden Daten auf verschiedene Arten erhoben und auch wenn Daten durch ihre Bezeichnung gleich erscheinen, können sich ihr Umfang und ihre Definition unterscheiden. IMPACT-2 hat deshalb einen Prozess entwickelt, nach dem generische Szenarien für vier Marktsegmente (HGV, Regionalverkehr, Metro und Güterverkehr) erstellt werden. Dabei werden nicht nur verschiedene DQ, sondern auch Expertenwissen aus unterschiedlichen Bereichen des Eisenbahnwesens aus weiten Teilen Europas genutzt. Durch diesen umfassenden Prozess sind die Ist-Szenarien und somit die Bewertungen adaptierbar für viele europäische Länder. Aufgrund der Struktur der in IMPACT-2 vertretenen Partner sowie der S2R-Mitglieder stammt der Großteil der Daten aus dem Norden und Westen Europas. Die Übertragbarkeit der Szenarien auf alle EU-Mitgliedsstaaten kann also nicht garantiert werden. Der gewählte Ansatz zur Vereinheitlichung europäischer Bahndaten kann jedoch für DQ zusätzlicher Länder erweitert werden. ■

LITERATUR | LITERATURE

- [1] Shift2Rail, 2015. Shift2Rail Multi-Annual Action Plan (MAAP). Shift2Rail, Brussels
- [2] COUNCIL REGULATION (EU) No 642/2014 of 16 June 2014, Article 2 s. 2a-c
- [3] IMPACT-1, 2017. D3.3 Use cases for SPDs. Shift2Rail, Brussels
- [4] IMPACT-1, 2018. D4.1 Reference Scenario. Shift2Rail, Brussels
- [5] IMPACT-2, 2019. D4.3 Reviewed quantitative KPI model. Shift2Rail, Brussels
- [6] Platform of Railway Infrastructure Managers in Europe, 2019. Key Performance Indicators for performance benchmarking. PRIME, Brussels
- [7] IMPACT-1, 2018. D4.2 Subsystem structure and sublevel KPIs. Shift2Rail, Brussels

Through this process, scenarios could be developed which allow being adaptable for most European countries.

Because of the partner structure within IMPACT-2 and the members of S2R, most of the available data are from Northern and Western European countries. Hence, the scenarios developed are not fully applicable for the whole European Union. Nevertheless, the used approach for unifying railway data can be extended to include data from further EU countries if these are available. ■

AUTOREN | AUTHORS

Alessa Eckert, M.Sc.

Wiss. Mitarbeiterin – Bewertung des Verkehrs /
Scientific staff – Evaluation of Transportation
Deutsches Zentrum für Luft- und Raumfahrt e.V. / German Aerospace Center
Institut für Verkehrssystemtechnik / Institute of Transportation Systems
Anschrift / Address: Rutherfordstr. 2, D-12489 Berlin
E-Mail: alessa.eckert@dlr.de

Svenja Hainz, M.Sc.

Wiss. Mitarbeiterin – Bewertung des Verkehrs /
Scientific staff – Evaluation of Transportation
Deutsches Zentrum für Luft- und Raumfahrt e.V. / German Aerospace Center
Institut für Verkehrssystemtechnik / Institute of Transportation Systems
Anschrift / Address: Lilienthalplatz 7, D-38108 Braunschweig
E-Mail: svenja.hainz@dlr.de



Sie wollen 1 Jahr im Gespräch bleiben?

Dann schalten Sie im EIK – EISENBAHN INGENIEUR KOMPENDIUM!

Wir schenken Ihnen mindestens einen Eintrag!

Bei der Buchung einer 1/1-Seite erhalten Sie zwei kostenlose Firmeneinträge.

Sie suchen die besondere Werbeform?

Dann ist unser Business-Profil für Sie genau das Richtige!
Stellen Sie Ihr Unternehmen detailliert vor und sichern Sie sich zusätzlich zwei kostenlose Einträge!

Gern sende ich Ihnen den aktuellen Themenplan zu.



Ihr Ansprechpartner: Tim Feindt ■ tim.feindt@dvvmedia.com ■ Telefon +49 40 237 14 220

